DESCRIPTION

Method and Apparatus for Face Description and Recognition using High-order Eigencomponents

Technical Field

The present invention relates to method and apparatus for face description and recognition using High-order Eigencomponents. The invention can be used in face description and recognition for content-based image retrieval, human face identification and verification for bank, security system and videophone; surveillance and tracking; digital library and Internet multimedia database.

Background Art

Human face perception is an active area in the computer vision community. Face recognition will play an important role in multimedia database search and many other applications. In recent years considerable progress has been made on the problems of face detection and recognition. Different techniques were proposed. Among these, neural networks, elastic template matching, Karhunen-Loeve expansion, algebraic moments and isodensity lines are typical methods.

Among these methods, principal component analysis

(PCA) or Karhunen-Loeve expansion is an important branch. Eigenface method is derived from PCA and it is convenient to be computed and has consistent accuracy in identification. Prior work showed that the PCA approach dissociates spontaneously between different types of information. The eigenvectors with large eigenvalues capture information that is common to subsets of faces and eigenvectors with small eigenvalues capture information specific to individual face. The studies show that only the information contained by the eigenvectors with large eigenvalues can be generalized to new faces that are not trained.

The advantage of eigenface method is that eigenvectors with large eigenvalues convey information relative to the basic shape and structure of the faces. That means features extracted from eigenvectors with large eigenvalues can be used to describe major characteristics of human faces. However, this is also the weakness of PCA. If we only consider the features extracted from eigenvectors with large eigenvalues, we cannot get the details of faces which corresponds to individual face. If these details of individual face can be described with the common features of human faces, the description of human faces can be more accurate.

One drawback of eigenface method is that the

contributions of all facial component areas are same. Instead of evenly distributed on the whole face, the identity information mainly locates at the certain facial areas, such as eyes, eyebrows, nose, mouth and outline. The check area contains less identity information and is relatively sensitive to lighting condition changes and facial expression changes. If the identity significances of facial components can be used, the recognition of human face can be more accurate.

Disclosure of Invention

Eigenface method is effective to extract common face characteristics such as shape and structure. In order to get details of faces that are lost when eigenvectors with small eigenvalues are truncated, the reconstructed faces with the features from eigenvectors with large eigenvalues should be obtained. With the reconstructed face images, the residue images between original images and reconstructed matrices can be obtained. These residue faces can be looked as high-passed face images, which still contain rich detailed information for individual face. In order to describe these residue faces, eigenface method can be used on these residue faces again. The obtained eigenvectors with large eigenvalues will reveal the common characteristics of residue faces. With this method, high-order eigenvectors with large eigenvalues can be obtained to extract

corresponding features. The combination of these features from different order eigenfaces can be used to describe faces effectively.

Similarly, first order and higher order principal components (eigencomponents) of facial components can be obtained to describe the characteristics of the corresponding facial areas. The combination of these features from different order eigencomponents can be used to describe individual facial components efficiently. Finally, human faces can be represented by a combination of different order eigencomponents with different attention weights. For different application fields, different components should have different functionalities (its strongerness or weakness). Different weights should be assigned for that component.

The present invention provides a method to interpret human faces which can be used for image retrieval (query by face example), person identification and verification, surveillance and tracking, and other face recognition applications. In order to describe face characteristics, the concept of high-order eigencomponents is proposed according to our observation and derivation. At first, all face images are normalized to a standard size. Then the vertical location of eyes is calculated and the face is shifted to a suitable place. When all these pre-processing

procedures are finished, the eigencomponents and high-order eigencomponents can be derived from a set of training face images. In order to query a face image in a face database, the features of the image projected with eigencomponents and high-order eigencomponents can be calculated with the selected eigencomponents and high-order eigencomponents. The combination of these features can be used to describe faces. With this description, Euclidean distance can be used for similarity measurement. In order to improve the similarity accuracy, the features should be weighted.

Brief Description of Drawings

Figure 1 shows a flowchart of the procedure for computing first -order feature $W^{(1)}$.

Figure 2 shows a flowchart of the procedures for computing i^{th}-order eigencomponents $U^{(i)}$ and the corresponding transform matrix U_i.

Figure 3 shows a flowchart for the training mode operation.

Figure 4 shows a flowchart for the test mode operation.

Best Mode for Carrying Out the Invention

The present invention gives a method to extract higher order eigencomponent features and represent a face by

combining different order component features.

With the normalized face images, the eigencomponents and high-order eigencomponents can be obtained as follows.

First use a preset block for the normalized face images to get the facial component (such as eyes, eyebrows, nose, mouth and outline);

Consider a facial component Φ_i which is a one-dimensional vector of raster-scanned facial component, define Ψ as the average component:

$$\Psi = \frac{1}{M}\sum_{i=1}^{M}\Phi_i \qquad (1)$$

Every facial component differs from the average component by a vector

$$\Gamma_i^{(1)} = \Phi_i - \Psi . \qquad (2)$$

The covariance matrix of the data is thus defined as:

$$Q = A^{(1)}A^{(1)^T} \qquad (3)$$

where

$$A^{(1)} = \left[\Gamma_1^{(1)}\Gamma_2^{(1)}\ldots\Gamma_M^{(1)}\right]. \qquad (4)$$

Note that Q has dimension $wh \times wh$ where w is the width of the component and h is the height. The size of this matrix is enormous, but since we only sum up a finite number of component vectors M, the rank of this matrix can not exceed M-1. We note that if $v_i^{(1)}$ is the eigenvector of $A^{(1)^T}A^{(1)}$

(i=1,2,...,M), then $A^{(1)T}A^{(1)}v_i^{(1)} = \lambda_i^{(1)}v_i^{(1)}$ where $\lambda_i^{(1)}$ are the eigenvalue of $A^{(1)T}A^{(1)}$, then $A^{(1)T}v_i^{(1)}$ are the eigenvectors of $A^{(1)}A^{(1)T}$ as we see by multiplying on the left by $A^{(1)}$ in the previous equation:

$$A^{(1)}A^{(1)T}A^{(1)}v_i^{(1)} = A^{(1)}\lambda_i^{(1)}v_i^{(1)} = \lambda_i^{(1)}A^{(1)}v_i^{(1)} \qquad (5)$$

Thus, eigenvector $v_i^{(1)}$ and eigenvalue $\lambda_i^{(1)}$ are obtained by the following equation.

$$\left[v_i^{(1)}, \lambda_i^{(1)}\right] = \mathrm{eig}(A^{(1)T}A^{(1)}, i), \quad i = 1,\ldots,M_1 \qquad (6)$$

But $A^{(1)T}A^{(1)}$is only of size $M \times M$. So defining $u_i^{(1)}$the eigenvector of $A^{(1)}A^{(1)T}$we have

$$u_i^{(1)} = A^{(1)}v_i^{(1)} = \sum_{k=1}^{M} v_{ik}^{(1)}\Gamma_k^{(1)} \qquad (7)$$

The eigenvalue $\lambda_i^{(1)}$ is the variance along the new coordinate space spanned by eigenvectors $u_i^{(1)}$. From here on we assume that the order of *i* is such that the eigenvalues $\lambda_i^{(1)}$are decreasing. The eigenvalues are decreasing in exponential fashion. Therefore we can project a facial component $\Gamma^{(1)}$onto only $M_1 \ll M$ dimensions by computing $W^{(1)} = \{w_k^{(1)}\}$ where $w_k^{(1)} = u_k^{(1)T}\Gamma^{(1)}$ and $1 \le k \le M_1$. $w_k^{(1)}$ is the *k*-th coordinate of $\Gamma^{(1)}$in the new coordinate system. In

this context, $W^{(1)}$ is called first -order component features. The vectors $u_k^{(1)}$ are actually images, and are called first-order eigencomponents.

Let

$$U^{(1)} = [u_1^{(1)} u_2^{(1)} \ldots u_{M_1}^{(1)}], \qquad (8)$$

then

$$W^{(1)} = U^{(1)^T} \Gamma^{(1)}. \qquad (9)$$

Since $U^{(1)}$ is an $M_1 \times P$ matrix, we cannot get its inverse. However, we can use its pseudo inverse to approximate its inverse. Let $U^{(1)^+}$ be the pseudo-inverse of $U^{(1)^T}$ as indicated below,

$$U^{(1)+} = \text{Pseudo} - \text{Inv}(U^{(1)}) \qquad (10)$$

then

$$\hat{\Gamma}^{(1)} = U^{(1)^+} W^{(1)} \qquad (11)$$

where $\hat{\Gamma}^{(1)}$ is the reconstructed matrix from $W^{(1)}$ and $U^{(1)}$.

Then, the following equation is carried out to obtain a residue component $\Gamma_i^{(2)}$.

$$\Gamma_i^{(2)} = \Gamma_i^{(1)} - \hat{\Gamma}^{(1)} \qquad (12)$$

Since a residue facial component vector still contains rich information for the individual component, facial component features should be extracted from the residue component

again. Let $A^{(2)} = \left[\Gamma_1^{(2)}\Gamma_2^{(2)}...\Gamma_M^{(2)}\right]$, $\lambda_i^{(2)}$ be the eigenvalues of $A^{(2)T}A^{(2)}$ and $v_i^{(2)}$ be the corresponding eigenvectors of $A^{(2)T}A^{(2)}$. Then $A^{(2)T}A^{(2)}v_i^{(2)} = \lambda_i^{(2)}v_i^{(2)}$. Based on above discussion, the eigenvectors of $A^{(2)}A^{(2)T}$ are $u_i^{(2)} = A^{(2)}v_i^{(2)}$. Therefore we can project a residue component $\Gamma^{(2)}$ onto only $M_2 << M$ dimensions by computing $W^{(2)} = \left\{w_k^{(2)}\right\}$, where

$$w_k^{(2)} = u_k^{(2)T}\Gamma^{(2)} \qquad (13)$$

and $1 \le k \le M_2$. Since $u_k^{(2)}$ are the eigenvectors of the residue component, we call $u_k^{(2)}$ the second -order eigencomponents and $w_k^{(2)}$ the second -order component features.

Let

$$U^{(2)} = \left[u_1^{(2)}u_2^{(2)}...u_{M_2}^{(2)}\right], \qquad (14)$$

Eq.(13) can be written as

$$\begin{aligned} W^{(2)} &= U^{(2)T}\Gamma^{(2)} \\ &= U^{(2)T}\left(\Gamma^{(1)} - \hat{\Gamma}^{(1)}\right) \\ &= U^{(2)T}\left(\Gamma^{(1)} - U^{(1)+}W^{(1)}\right) \\ &= U^{(2)T}\left(\Gamma^{(1)} - U^{(1)+}U^{(1)T}\Gamma^{(1)}\right) \\ &= \left(U^{(2)T} - U^{(2)T}U^{(1)+}U^{(1)T}\right)\Gamma^{(1)} \end{aligned} \qquad (15)$$

Let

$$U_2 = \left(U^{(2)T} - U^{(2)T}U^{(1)+}U^{(1)T}\right)^T, \qquad (16)$$

we have

$$W^{(2)} = U_2^T\Gamma^{(1)}. \qquad (17)$$

Since U_2 is a constant transform matrix and it is just calculated once, it will not affect the efficiency of

computation. The facial component can be described with

$$\Omega(\Phi)=\left[w_1^{(1)},w_2^{(1)},\ldots,w_{M'_1}^{(1)},w_1^{(2)},w_2^{(2)},\ldots,w_{M_2}^{(2)}\right]^T, \tag{18}$$

where $1\leq M_1'\leq M_1$. The computational burden does not increase in computing $\Omega(\Phi)$ compared with only computing component features from eigencomponents U.

The residue components are called second -order residue components and the original components are called first -order residue components.

With the same method, we also can derive third-order, fourth-order ,..., and n^{th}-order eigencomponents. By projecting the residue components of corresponding order, we can get third-order, fourth-order ,..., n^{th}-order component features. With these high-order component features, the similarity of components can be defined as the weighted Euclidean distance between the projections. Fig. 2 illustrates the procedure to compute the i^{th}-order eigencomponents $U^{(i)}$ and the corresponding transform matrix U_i. In the figure, Pseudo_Inv(B) is the function to calculate the pseudo-inverse of matrix B.

The measure of dissimilarity of two faces H_1 and H_2 is defined as a combined distance between various facial component features generated from the projections of eigencomponents (i.e. eigeneyes, eigeneyebrows,

eigennoses, eigenmouths and eigenoutlines) and eigenfaces.

$$
\begin{aligned}
D(H_1,H_2)= &\sum_{i=1}^{M'^{eye}_1} a_1^{eye}\left\|w_i^{(1)}\left(\Phi_1^{eye}\right)-w_i^{(1)}\left(\Phi_2^{eye}\right)\right\|+\sum_{j=1}^{M^{eye}_2} a_2^{eye}\left\|w_j^{(2)}\left(\Phi_1^{eye}\right)-w_j^{(2)}\left(\Phi_2^{eye}\right)\right\| \\
&+\sum_{i=1}^{M'^{eyebrow}_1} a_1^{eyebrow}\left\|w_i^{(1)}\left(\Phi_1^{eyebrow}\right)-w_i^{(1)}\left(\Phi_2^{eyebrow}\right)\right\|+\sum_{j=1}^{M^{eyebrow}_2} a_2^{eyebrow}\left\|w_j^{(2)}\left(\Phi_1^{eyebrow}\right)-w_j^{(2)}\left(\Phi_2^{eyebrow}\right)\right\| \\
&+\sum_{i=1}^{M'^{nose}_1} a_1^{nose}\left\|w_i^{(1)}\left(\Phi_1^{nose}\right)-w_i^{(1)}\left(\Phi_2^{nose}\right)\right\|+\sum_{j=1}^{M^{nose}_2} a_2^{nose}\left\|w_j^{(2)}\left(\Phi_1^{nose}\right)-w_j^{(2)}\left(\Phi_2^{nose}\right)\right\| \\
&+\sum_{i=1}^{M'^{mouth}_1} a_1^{mouth}\left\|w_i^{(1)}\left(\Phi_1^{mouth}\right)-w_i^{(1)}\left(\Phi_2^{mouth}\right)\right\|+\sum_{j=1}^{M^{mouth}_2} a_2^{mouth}\left\|w_j^{(2)}\left(\Phi_1^{mouth}\right)-w_j^{(2)}\left(\Phi_2^{mouth}\right)\right\| \\
&+\sum_{i=1}^{M'^{outline}_1} a_1^{outline}\left\|w_i^{(1)}\left(\Phi_1^{outline}\right)-w_i^{(1)}\left(\Phi_2^{outline}\right)\right\|+\sum_{j=1}^{M^{outline}_2} a_2^{outline}\left\|w_j^{(2)}\left(\Phi_1^{outline}\right)-w_j^{(2)}\left(\Phi_2^{outline}\right)\right\| \\
&+\sum_{i=1}^{M'_1} a_1\left\|w_i^{(1)}\left(H_1\right)-w_i^{(1)}\left(H_2\right)\right\|+\sum_{j=1}^{M_2} a_2\left\|w_j^{(2)}\left(H_1\right)-w_j^{(2)}\left(H_2\right)\right\|
\end{aligned}
$$

If $a_1=0$, the similarity of face images will be measured only with second -order features.

With the method mentioned above, the method for describing a face image can be considered as follows:

1) Scanning the said facial component by using a raster scan starting at the top-left corner of the component window and finishing at the bottom-right corner of the component window into a single dimensional array of pixels;

2) Subtracting the said single dimensional array of pixels with the average component;

3) Multiplying the said subtracted single dimensional array of pixels with the said first-order and high-order eigencomponents;

4) Using the resulting component features as description of the face;

5) Coding the said features into a coded representation.

With this method, the human faces can be described effectively and efficiently in space with different attention weights corresponding to the significance of identity information of human face.

Next, an overall operation of the facial feature extraction according to the present invention is described with reference to Figs. 3 and 4. The operation includes a training mode operation (steps #22-#31) as shown in Fig. 4, and a test mode operation (steps #32-#42) as shown in Fig. 3. The training mode operation is carried out first to learn and accumulate many face samples and to obtain coefficients of a first order average.

The training mode starts from step #22 and continues to step #31. The training mode is provided to generate various parameters to be used in the test mode.

In step #22, a plurality of sample face images are input.

In step #23, each sample face image is divided into a plurality of facial parts, such as right eye, left eye, right eyebrow, left eyebrow, nose, mouth, facial configuration, and each part is analyzed to obtain basic facial component Φ_i. The facial component Φ_i can be weighted according to the significance of identity of human face.

In step #24, the facial components of the same part, such as a nose part, are collected from a plurality of sample face images. The facial component of nose part is referred to as a nose component. The collected facial components

of each part are averaged to obtain a first order average facial component Ψ by using equation (1). The same operation is carried out to obtain the first order average facial components for different facial components. The first order average facial component Ψ is used in the test mode operation in step #34.

In step #25, the facial component of, for example, nose from each sample face image is subtracted by the first order average facial component Ψ of nose to obtain a vector of nose given by equation (2). The same operation is carried out for each of different facial components.

Steps #22 to #25 taken together is called an analyzing step for analyzing the training face images.

In step #26, equations (4), (5), (6), (7) and (8) are carried out to obtain the first order eigencomponents $U^{(1)}$. The first order eigencomponents $U^{(1)}$ is used in the test mode operation in step #35.

In step #27, an inverse matrix $U^{(1)+}$ is generated by using equation (10). The inverse matrix $U^{(1)+}$ is nearly equal to $U^{(1)T}$. The inverse matrix $U^{(1)+}$ is used in the test mode operation in step #37.

In step #28, the inverse matrix $U^{(1)+}$ is used for obtaining a difference of the facial component using equation (11). Thus, a difference of facial component with respect to the basic facial component of the original facial images as collected in step #22 is obtained. The data produced in step #28 is referred to as a difference facial component.

In step #29, a difference between the difference facial component and the first order average facial component is obtained.

Steps #27 to #29 taken together is called an analyzing step for analyzing the first -order eigencomponent.

In step #30, using the difference obtained in step #29, a second order K-L coefficient $U^{(2)}$ (which is also called second order eigencomponent) is calculated using equations (4), (6), (7), (14) and (16). The second order K-L coefficient $U^{(2)}$ is used in the test mode in K-L conversion (step #40).

The test mode starts from step #32 and continues to step #42. The test mode is provided to generate first order component feature $W^{(1)}$ and second order component feature $W^{(2)}$

In step #32, a face to be tested is input.

In step #33, the input face image is divided into a plurality of facial parts, such as right eye, left eye, right eyebrow, left eyebrow, nose, mouth, facial configuration, and each part is analyzed to obtain basic facial component Φ_i. The facial component Φ_i can be weighted according to the significance of identity of human face.

In step #34, the basic facial component is subtracted by the first order average facial component Ψ to obtain a first difference $\Gamma^{(1)}$ (also called a first order residue component). The first difference $\Gamma^{(1)}$ is applied to step #35 and to step #39.

Steps #32 to #34 taken together is called an analyzing step for analyzing the test face image.

In step #35, using the first difference $\Gamma^{(1)}$ and the first order eigencomponents $U^{(1)}$, K-L conversion is carried out by the use of equation (9) to obtain the first order component feature $W^{(1)}$.

In step #36, the first order component feature $W^{(1)}$ is produced. The first order component feature $W^{(1)}$ represents the feature of the test face input in step #32. The first order component feature $W^{(1)}$ can be used as information

representing the test face, but has a relatively large size of data. A further calculation is carried out to reduce the data size.

In step #37, using the equation (11), K-L inverse conversion is carried out to generate a reconstructed matrix $\hat{\Gamma}^{(1)}$.

In step #38, the reconstructed matrix $\hat{\Gamma}^{(1)}$ is produced.

In step #39, the equation (12) is carried out to produce a difference between the first difference $\Gamma^{(1)}$ and the reconstructed matrix $\hat{\Gamma}^{(1)}$ to obtain a second difference $\Gamma^{(2)}$, which is generally referred to as a second residue component $\Gamma_i^{(2)}$.

Steps #37 to #39 taken together is called an analyzing step for analyzing the first -order component feature.

In step #40, using the second difference $\Gamma^{(2)}$ and the second order K-L coefficient $U^{(2)}$, K-L conversion is carried out by equation (17) to generate the second order component feature $W^{(2)}$.

In step #41, the second order component feature $W^{(2)}$ is produced. The second order component feature $W^{(2)}$ has information representing the test face as entered in the test mode.

It is to be noted that the flowcharts shown in Figs. 3 and 4 can be arranged by a computer connected with a camera for capturing the sample face images and the test face image. It is possible to prepare two sets of arrangements, one set is for the training mode operation, and another set for the test mode operation. Each set includes a computer and a camera. The set for the training mode operation is programmed to carry out steps #22-#30, and the set for the test mode operation is programmed to carry out steps #32-#42. In the set for the test mode operation, a memory is provided to previously store

information obtained by the set for the training mode operation, such as the first order average facial component Ψ, the first order eigencomponents $U^{(1)}$, the inverse matrix $U^{(1)+}$, and the second order eigencomponent $U^{(2)}$.

This invention is very effective for describing human faces using component-based features. Since the high-order eigencomponents can be calculated only once with the training components, the high-order component features can be obtained as efficient as first-order component features. However, since detailed regional identity information can be revealed with high-order component features, the combination of first-order component features and high-order component features of eyes, eyebrows, nose, mouth and outline with different attention weights has better face description capability compared with the first -order eigenface features or combined first -order and high-order eigenface features.

This invention is very effective and efficient in describe human faces which can be used in internet multimedia database retrieval, video editing, digital library, surveillance and tracking, and other applications using face recognition and verification broadly.